

**IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF VIRGINIA
ALEXANDRIA DIVISION**

ROSY GIRON DE REYES, *et al.*,

Plaintiffs,

v.

WAPLES MOBILE HOME PARK
LIMITED PARTNERSHIP, *et al.*,

Defendants.

Civil No.: 1:16cv563-TSE-TCB

**MEMORANDUM IN SUPPORT OF
DEFENDANTS' MOTION TO STRIKE THE OPINIONS AND TESTIMONY OF
PROFESSOR WILLIAM A.V. CLARK**

TABLE OF CONTENTS

	Page
INTRODUCTION	1
LEGAL STANDARD.....	3
PROFESSOR CLARK’S OPINIONS	5
ARGUMENT	9
I. PROFESSOR CLARK’S METHODOLOGY IS NOT RELIABLE BECAUSE HE DID NOT DETERMINE A MARGIN OF ERROR OR CONFIDENCE LEVEL FOR HIS STATISTICAL ESTIMATES OF UNDOCUMENTED HISPANICS.....	9
II. PROFESSOR CLARK’S FAILURE TO DETERMINE AN MOE FOR UNDOCUMENTED ASIANS LIKEWISE RENDERS HIS OPINION THAT UNDOCUMENTED HISPANICS ARE IMPACTED MORE THAN UNDOCUMENTED ASIANS COMPLETELY UNRELIABLE	15
III. PROFESSOR CLARK INCORRECTLY APPLIED CMS’S 31.4% ESTIMATE FOR UNDOCUMENTED HISPANICS IN THE PUMA TO THE MUCH SMALLER CENSUS TRACT BASED ON HIS <i>IPSE DIXIT</i> ASSUMPTION THAT THE DEMOGRAPHICS WOULD BE THE SAME.....	18
IV. EXCLUSION IS ALSO WARRANTED UNDER RULE 403	19
CONCLUSION.....	20

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Anderson v. Westinghouse Savannah River Co.</i> , 406 F.3d 248 (4th Cir. 2005)	4, 15
<i>Cooper v. Smith & Nephew, Inc.</i> , 259 F.3d 194 (4th Cir. 2001)	18
<i>Crawford v. Newport News Indus. Corp.</i> , No. 4:14-CV-130, 2017 WL 3222547 (E.D. Va. July 28, 2017).....	15
<i>Daubert v. Merrell Dow Pharmaceuticals</i> , 509 U.S. 579 (1993).....	<i>passim</i>
<i>EEOC v. Freeman</i> , 961 F. Supp. 2d 783 (D. Md. 2013), <i>aff'd in part sub nom. E.E.O.C. v. Freeman</i> , 778 F.3d 463 (4th Cir. 2015)	<i>passim</i>
<i>Evenwel v. Abbott</i> , 2015 WL 5675832 (U.S., 2015).....	13, 18
<i>Gen. Electric Co. v. Joiner</i> , 522 U.S. 136 (1997).....	5
<i>Harper v. Computer Assocs. Int'l, Inc.</i> , No. 97 C 5954, 1998 WL 547297 (N.D. Ill. Aug. 27, 1998).....	10
<i>Hutchinson v. Hamlet</i> , No. C 02-974 JSW (PR), 2006 WL 1439784 (N.D. Cal. May 23, 2006)	11
<i>Kumho Tire Co. v. Carmichael</i> , 526 U.S. 137 (1999).....	4, 18
<i>Lucas v. Fifth Third Mortg. Co.</i> , No. CIV.A. 3:13-2194, 2013 WL 1437738 (S.D.W. Va. Apr. 9, 2013).....	10
<i>Montes v. City of Yakima</i> , 40 F. Supp. 3d 1377 (E.D. Wash. 2014).....	11
<i>Nease v. Ford Motor Co.</i> , 848 F.3d 219 (4th Cir. 2017)	11
<i>Oglesby v. Gen. Motors Corp.</i> , 190 F.3d 244 (4th Cir. 1999)	4

<i>Payne v. Tri-State CareFlight, LLC</i> , 332 F.R.D. 611 (D.N.M. 2019), <i>leave to appeal denied</i> , No. 19-702, 2019 WL 8329300 (10th Cir. Dec. 3, 2019)	10
<i>Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc.</i> , 576 U.S. 519 (2015)	1
<i>Valador, Inc. v. HTC Corp.</i> , 242 F. Supp. 3d 448 (E.D. Va. 2017)	4
<i>Watson v. Fort Worth Bank & Tr.</i> , 487 U.S. 977 (1988)	4, 9
<i>Weisgram v. Marley Co.</i> , 528 U.S. 440 (2000)	4
<i>Westberry v. Gislaved Gummi AB</i> , 178 F.3d 257 (4th Cir. 1999)	4
Rules	
Rule 403	3, 5, 19, 20
Rule 702	3, 15, 20
Other Authorities	
Reference Guide on Statistics, Reference Manual on Scientific Evidence (3d ed. 2011) 2011 WL 7724256, available at https://www.fjc.gov/content/reference-manual-scientific-evidence-third-edition-1	10, 11, 13, 14
<i>A Basic Explanation of Confidence Intervals</i> , U.S. Census, available at https://www.census.gov/programs-surveys/saipe/guidance/confidence-intervals.html (last visited Sept. 18, 2020)	10

INTRODUCTION

Plaintiffs’ claim of an alleged disparate impact on Hispanics by Defendants’ Policy is based upon the fundamentally flawed opinions of their statistical expert, Professor Clark, that fail, by Professor Clark’s own admission, to adhere to basic principles of statistical analysis. Chief among these is the failure of Professor Clark to provide a margin of error (“MOE”) for his statistical conclusions – the basic reliability test for any statistical estimate. And he applies population statistics between dramatically different size geographical areas with no rhyme or reason – again eschewing basic statistical principles. Accordingly, his opinions must be rejected.

Professor Clark’s opinions are central to the prima facie case that Plaintiffs must establish in order to prove their disparate impact claim under the three step burden shifting scheme set forth in *Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc.*, 576 U.S. 519 (2015). Pursuant to *Inclusive Communities*, “[a] plaintiff who fails to... produce statistical evidence demonstrating a causal connection cannot make out a prima facie case of disparate impact.” 576 U.S. at 543. Plaintiffs’ disparate impact claim is predicated on allegations of statistics purportedly demonstrating a disparate impact on Hispanics resulting from Defendants’ Policy, Compl. ¶¶ 58-63, and Plaintiffs identified Professor Clark to provide evidence supporting those statistical allegations.

Professor Clark’s opinions of disparate impact necessarily hinge on his estimates of the undocumented Hispanic population in the census tract that encompass Waples mobile home park (the “Park”).¹ To reach his opinions, Professor Clark relied on data from the Center of Migration

¹ Professor Clark concluded that Defendants’ Policy has a disparate impact on Latinos based on his analysis of statistics in two census defined areas that encompass the Park: A larger “Public Use Microdata Area” (“PUMA”), with approximately 160,000 persons, and a dramatically smaller area within the PUMA, Census Tract 4406, with approximately 3,300 persons.

Studies (“CMS”) for “all calculations”. However, CMS’s estimates suffer from a fundamental reliability problem. CMS does not publish an MOE for any of its estimates for the undocumented population below its estimate of the *national* undocumented population. As Professor Clark admitted, with any statistical estimate drawn from samples, an MOE is necessary to account for the error in the result because a sample is not a true count. CMS has no MOE below the national level because it is difficult to measure a population that does not want to be found. Professor Clark, however, never calculated an MOE for CMS’s estimates that he relied upon for his opinions. Instead, recognizing that an MOE is a necessary measure of an estimate’s reliability, he attempted to apply an MOE taken from inapposite statistics – the American Community Survey’s (“ACS”)² estimate of the *total* Hispanic population in Tract 4406 – and apply this MOE to his estimate of the undocumented Hispanics in Tract 4406. The MOE is inapposite because it does not take into account the admitted difficulties of estimating the undocumented population.

Additionally, because CMS does not publish an undocumented population estimate at a level smaller than the PUMA level, Professor Clark also faced a problem of how to determine the undocumented population at the much smaller census tract level. Attempting to resolve this problem, he simply assumed that CMS’s estimate of the undocumented population at the PUMA level (that 31.4% of the Hispanic population was undocumented) would apply to the dramatically smaller sample size of Tract 4406—even while admitting that he did nothing to determine whether the demographics of Tract 4406 are the same or similar to the demographics of the much larger

² The American Community Survey is the largest household survey in the United States. Dkt. 248-3, Ex. 3(1) at “Basis and Reasons for Opinions”. It is a nation-wide survey that samples approximately 3.5 million addresses. Understanding and Using ACS Data; What All Data Users Need to Know at 1, Chapter 1: Understanding the ACS: The Basics, U.S. Census Bureau, available at <https://www.census.gov/programs-surveys/acs/guidance/handbooks/general.html>. The ACS data relied upon by Professor Clark does not include estimates of the undocumented population.

PUMA, despite the fact that Tract 4406 is dramatically smaller and variability generally increases with smaller sample sizes.

Compounding this problem, Professor Clark opined that Latinos would suffer a disparate impact from Defendants' Policy, but Professor Clark initially failed to compare undocumented Hispanics to undocumented Asians in the relevant areas. In his rebuttal report, Professor Clark made this comparison for the first time, and in so doing, drastically reduced his estimate of how much more likely Hispanics would be impacted when compared to Asians. But, Professor Clark's belated comparison suffered from the same fundamental flaws as his initial estimates regarding undocumented Hispanics – he failed to determine an MOE for his estimates of undocumented Asians rendering his estimates of undocumented Asians and his comparison wholly unreliable.

This Court should therefore exercise its gate-keeping role pursuant to *Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579, 595 (1993) and exclude Professor Clark's unreliable opinions. Finally, Rule 403 separately counsels in favor of excluding his opinions because they lack probative value and are unfairly prejudicial.

LEGAL STANDARD

Under Rule 702, “[a] witness who is qualified as an expert by knowledge, skill, experience, training, or education” may only testify and provide an expert opinion on four conditions: (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue; (b) the testimony is based on sufficient facts or data; (c) the testimony is the product of reliable principles and methods; and (d) the expert has reliably applied the principles and methods to the facts of the case. Fed. R. Evid. 702. This Court's “gatekeeping role” with respect to Rule 702 is vital given how “expert evidence can be both powerful and quite misleading.” *Daubert*, 509 U.S. at 595.

As this Court has explained, “given the potential persuasiveness of expert testimony,

proffered evidence that has a greater potential to mislead than to enlighten should be excluded.” *Valador, Inc. v. HTC Corp.*, 242 F. Supp. 3d 448, 456 (E.D. Va. 2017) (quoting *Westberry v. Gislaved Gummi AB*, 178 F.3d 257, 261 (4th Cir. 1999)). Thus, the Court must “ensur[e] that an expert’s testimony both rests on a reliable foundation and is relevant to the task at hand.” *Daubert*, 509 U.S. at 597. The “exacting standards of reliability” that the Supreme Court has adopted, *Weisgram v. Marley Co.*, 528 U.S. 440, 455 (2000), require “more than subjective belief or unsupported speculation[.]” *Daubert*, 509 U.S. at 590. The expert must “employ[] in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.” *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 152 (1999). To ensure as much, the Court should scrutinize the expert’s “principles and methodology.” *Daubert*, 509 U.S. at 595. As the Fourth Circuit put it: “A reliable expert opinion must be based on scientific, technical, or other specialized knowledge and not on belief or speculation, and inferences must be derived using scientific or other valid methods.” *Oglesby v. Gen. Motors Corp.*, 190 F.3d 244, 250 (4th Cir. 1999).

As for statistics in particular, “[e]xpert statistics are unreliable if they are based on ‘incomplete data sets and inadequate statistical techniques.’” *EEOC v. Freeman*, 961 F. Supp. 2d 783, 792 (D. Md. 2013), *aff’d in part sub nom. E.E.O.C. v. Freeman*, 778 F.3d 463 (4th Cir. 2015) (citing *Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977, 996–97 (1988)). Among other factors, the district court should “consider the rate of error of the methods employed by the expert, the existence and maintenance of standards used in the expert’s methods, and whether the expert’s methods have been generally accepted by his or her respective community.” *Anderson v. Westinghouse Savannah River Co.*, 406 F.3d 248, 261 (4th Cir. 2005). “[A] court should exclude expert statistical testimony when the data relied upon is connected to the expert’s opinion ‘only by the ipse dixit of the expert. A court may conclude that there is simply too great an analytical gap

between the data and the opinion proffered.” *Freeman*, 961 F. Supp. 2d at 793 (quoting *Gen. Electric Co. v. Joiner*, 522 U.S. 136, 146 (1997)).

Rule 403 separately permits the exclusion of relevant evidence “if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury.” *Daubert*, 509 U.S. at 595. “Expert evidence can be both powerful and quite misleading because of the difficulty in evaluating it. Because of this risk, the judge . . . exercises more control over experts than over lay witnesses.” *Id.* (quoting 138 F.R.D. 631, 632 (1991)).

PROFESSOR CLARK’S OPINIONS

In his October 28, 2016 opening report, Professor Clark opined that “[a]ny policy focused specifically on the sub-area defined by [Census] Tract 4406 will have a disproportionate impact on the Hispanic population.” **Exhibit 1** at 1 § 2. Thus, his analysis of the population of Tract 4406 is the crux of his opinion. In order to reach this opinion, he attempted to estimate, by using statistics, the number of undocumented Hispanics in the census tract where the Park is located, Census Tract 4406. *Id.* at 1 § 1. He primarily relied on census data from the ACS and data from the CMS. *Id.* at 2 § 3. He relied on ACS data for total Hispanic populations and total populations in a geographic area. According to the ACS data, Tract 4406 has a total population of 3,294. *Id.* at 3 § 5. To estimate the undocumented population, he used CMS data for “all calculations.” *Id.* at 4 § 5 (Subsection “The Undocumented Population”). He admitted that CMS does not estimate the number of undocumented aliens for an area as small as a census tract. **Exhibit 2** (Clark Dep. Tr.) at 64:4-7. Instead, the “smallest relevant geographic area for which there are [CMS] undocumented population estimates is the Fairfax County Central, Fairfax City, Burke PUMA (“Public Use Microdata Area”). Ex. 1 at 4-5 § 5.

According to ACS data, the Fairfax PUMA that encompasses the Park has a total population of 157,949 with a Hispanic population of 18,954. *Id.* at 5 § 5. According to CMS data,

CMS estimates that 31.4% of the total Hispanic population at the PUMA level (18,954), or 5,944, is undocumented. *Id.* This CMS estimate, however, lacks any MOE to evaluate its reliability: Although CMS provides an estimate of the undocumented population at the PUMA level, Professor Clark admitted that CMS has not provided *any* MOE for its estimates of the undocumented population at a level smaller than the *national* level. Ex. 2 at 34:6-19; *accord id.* at 82:3-9 (“CMS did not provide margins of error at the PUMA level.”). However, Professor Clark testified that an MOE is necessary because “when statisticians and demographers make estimates, they recognize that there is some – because it is not a count – error in the result.” *Id.* at 26:19-22; Ex. 1 § 5 § 5 (noting “of course” there are “variations in small area estimates (the Margin of Error).”)

Nonetheless, Professor Clark applied CMS’s 31.4% estimate for the PUMA to the much smaller population of Tract 4406³ by applying the 31.4% estimate to the total population of Hispanics in Tract 4406 (957) in order to conclude there are 301 undocumented Hispanics in Tract 4406. Ex. 1 at 5 § 5 (“The PUMA has a population of 157,949 and the Hispanic population is 18,954. Of the Hispanic population 5,944 (31.4%) are estimated to be undocumented (CMS database statistics from the PUMA data tool). We can apply that estimate to the 957 Hispanics in Tract 4406...”); Ex. 2 at 66:10-11 (“we apply the same proportion at the PUMA level down to the tract level.”).

Professor Clark applied ACS’s MOE for ACS’s estimate of *total* Hispanics in the census tract to *his* estimate of *undocumented* Hispanics in the census tract. Professor Clark states that the ACS MOE for the estimate of the *total Hispanic population* in Tract 4406 is 26%. Ex. 1 § 5. He

³ According to the statistics relied on by Professor Clark, Tract 4406 has a total population of 3,294 versus the much larger Fairfax PUMA’s population of 157,949. Ex. 1 at 4 (second table noting Tract 4406’s total population of 3,294); *id.* at 5 (“The PUMA has a population of 157,949.”).

then applied this 26% MOE to his estimate of 301 *undocumented* Hispanics in Tract 4406 to conclude a range of undocumented Hispanic population in Tract 4406 of 223-379. *Id.*; Ex. 2 at 37:21-38:4.⁴ He did not adjust the ACS MOE even though he admits that it is more difficult to estimate an undocumented population because they are “less willing, less wanting to be measured.” Ex. 2 at 33:5-16. According to Professor Clark, it is so difficult to estimate this population that CMS does not provide an MOE for their undocumented population estimates at any level smaller than the national level. *Id.* at 34:16-35:2.

Nonetheless, Professor Clark concludes that because undocumented Hispanics make up 31.4% of Tract 4406, with a 26% MOE, he concludes that Latinos are “nearly 7 times more likely to be undocumented than other groups and so 7 times more likely to be adversely affected by [Defendants’] policy.” *Id.* at 5 § 5. He subsequently recanted this conclusion in his reply report when he admitted that he failed to consider the Asian undocumented population in his analysis. **Exhibit 3** at 2 § 3 (“Dr. Weinberg is correct that I did not adjust for the Asian undocumented population.”). Instead, Professor Clark admitted in his rebuttal report that Latinos are not even twice as likely to be undocumented as Asians. *Id.* (“Latinos are nearly twice as likely to be undocumented compared to Asians and 20 times more likely to be undocumented than other groups.”).

Professor Clark followed the same methodology in calculating estimates of the undocumented Asian population. *Id.* at 2 § 3 (calculating the undocumented Asian population using the “same methodological procedure”). That is, he never determined an MOE for the estimates of the undocumented Asian population and merely adopted the ACS MOE for the total

⁴ Professor Clark never states what the ACS MOE is for the total Hispanic population at levels higher than the tract level or if he is adopting ACS’s MOE at any other level other than the tract level.

Asian population at the tract level for *his* estimate of the undocumented Asian population within the tract. He also applied the CMS estimate for the undocumented Asian population in the PUMA to the Tract as CMS does not publish estimates of the undocumented population at the tract level. And, crucially, Professor Clark never stated in his reports what confidence level he was using for any of his estimates, whether of the Hispanic or Asian populations, even though as he admits, “you can have margins of error at various confidence levels.” Ex. 2 at 28:18-22.

As far as the required method to calculate an MOE from CMS’s data, Professor Clark admitted that CMS’s only estimate of the undocumented population that has a published MOE is its estimate of the undocumented population at the national level, which has an MOE of 9%. Ex. 2 at 34:10-15. He admitted that as Census Bureau estimates go from the national level to smaller geographical areas, its margins of error increase. *Id.* at 32:9-13. And, he admitted that CMS does not provide an MOE for undocumented estimates smaller than the national level because “it’s difficult enough to try and get estimates of the undocumented population. Putting margins of error on this is difficult – a difficult process.” *Id.* at 34:16-35:2. Simply adopting the ACS MOE at the tract level, Professor Clark, however, never calculated a margin of error for CMS’s estimates of the undocumented population at levels smaller than the national level nor for his own estimate of the undocumented population in Tract 4406.

In response to Professor Clark’s opinions, Defendants identified an opposing statistical expert, Dr. Daniel H. Weinberg. Dr. Weinberg, who holds a Ph.D. and two master degrees in economics from Yale University, has spent 25 years as a senior management executive and research scientist at the U.S. Census Bureau. **Exhibit 4** at 7. He has served as a division director with responsibility for all housing surveys at the Census Bureau and as Assistant Director for the Decennial Census and the American Community Survey. *Id.*; *id.* at Ex. A. Dr. Weinberg criticized

Professor Clark's methodology as unreliable because, as he wrote in his report, CMS's only published MOE – 9% at the national level (11 million +/- 1 million) – results in an MOE of greater than 100% at the census tract level. *Id.* at 1 ¶ 1. Professor Clark, however, never calculated any MOE, much less an MOE for his statistical estimate that 31.4% of the Hispanic population in the Tract was undocumented.

Yet, when questioned about whether he should have calculated an appropriate MOE for his undocumented population estimates by using the only published MOE by CMS for the undocumented population – the 9% MOE at the national level – and then moving downward to smaller sample sizes and increasing the MOE to adjust for the variability of a smaller sample size, Professor Clark tellingly admitted that “[p]erhaps the margin of error should be larger.” Ex. 2 at 78:12-13.

ARGUMENT

I. PROFESSOR CLARK'S METHODOLOGY IS NOT RELIABLE BECAUSE HE DID NOT DETERMINE A MARGIN OF ERROR OR CONFIDENCE LEVEL FOR HIS STATISTICAL ESTIMATES OF UNDOCUMENTED HISPANICS

Professor Clark's opinions are not reliable because he does not provide an MOE regarding his estimate of undocumented Hispanics in the relevant PUMA or census tract. As Professor Clark himself admits, however, an MOE is an essential and basic requirement for any statistical analysis. In the absence of an MOE, there is no basis to determine if the statistical estimates are reliable because there is no basis to measure the rate of error of the estimate. Accordingly, the failure of Professor Clark to calculate an MOE for his estimates of the undocumented Hispanic population requires the exclusion of his opinions.

“Expert statistics are unreliable if they are based on... inadequate statistical techniques.” *Freeman*, 961 F. Supp. 2d at 792 (citing *Watson*, 487 U.S. at 996–97). If statistical estimates fail

to provide an MOE (also known as a confidence interval)⁵, they do not provide a measure to evaluate the reliability of the estimates. *Harper v. Computer Assocs. Int'l, Inc.*, No. 97 C 5954, 1998 WL 547297, at *5 (N.D. Ill. Aug. 27, 1998) (finding reliability of statistics “questionable” where expert failed “to address the standard deviation or margin of error.”); see *Lucas v. Fifth Third Mortg. Co.*, No. CIV.A. 3:13-2194, 2013 WL 1437738, at *3 (S.D.W. Va. Apr. 9, 2013) (“The Court similarly has no information regarding the confidence interval (also known as the margin of error) of the analysis.”) (remanding for lack of jurisdiction for failure to demonstrate CAFA’s statutory requirements).

When statistical estimates are derived from samples, necessarily because an entire population cannot be counted, MOEs need to be provided as a measure of the reliability of the estimate to account for error that arises because the sample may not accurately reflect the whole population.⁶ The U.S. Census Bureau explained this succinctly in its guide to using ACS data:

Because the ACS is based on a sample, rather than all housing units and people, ACS estimates have a degree of uncertainty associated with them, called sampling error. In general, the larger the sample, the smaller the level of sampling error. To help users understand the impact of sampling error on **data reliability**, the Census Bureau provides a “margin of error” for each published ACS estimate. The margin of error, combined with the ACS estimate, give users a range of values within which

⁵ “A confidence interval for a proportion estimate is also known as a ‘margin of error.’ It is the ‘plus or minus’ figure often displayed next to the proportion estimate in, e.g., public polling data.” *Payne v. Tri-State CareFlight, LLC*, 332 F.R.D. 611, 684 n.31 (D.N.M. 2019), *leave to appeal denied*, No. 19-702, 2019 WL 8329300 (10th Cir. Dec. 3, 2019); see also *A Basic Explanation of Confidence Intervals*, U.S. Census, available at <https://www.census.gov/programs-surveys/saipe/guidance/confidence-intervals.html> (last visited Sept. 18, 2020) (“A confidence interval is a range of values that describes the uncertainty surrounding an estimate.”).

⁶ “[S]ampling error. A sample is part of a population. When a sample is used to estimate a numerical characteristic of the population, the estimate is likely to differ from the population value because the sample is not a perfect microcosm of the whole.” Reference Guide on Statistics, Reference Manual on Scientific Evidence 296 (3d ed. 2011), 2011 WL 7724256. The Reference Manual on Scientific Evidence, Third Edition is also available at <https://www.fjc.gov/content/reference-manual-scientific-evidence-third-edition-1>.

the actual “real-world” value is likely to fall.⁷

In another chapter from the same handbook, the U.S. Census Bureau explained that “[a] margin of error (MOE) describes the precision of an ACS estimate at a given level of confidence.”⁸

According to the Reference Guide on Statistics published by the Federal Judicial Center, large MOEs may indicate that an estimate is “seriously wrong” or “less reliable.” Reference Guide on Statistics at 246, 2011 WL 7724256 (stating that if standard errors or confidence intervals are large, “the estimate may be seriously wrong”); *id.* at 247 (“[A] broader [confidence] interval indicates less precision.”); *id.* at 248 (“A high confidence level with a broad interval means very little.”);⁹ *Montes v. City of Yakima*, 40 F. Supp. 3d 1377, 1404 n.5 (E.D. Wash. 2014) (“[T]he narrower the confidence interval, the more reliable the estimate; the broader the confidence interval, the less reliable the estimate.”).

One of the *Daubert* factors for a court to consider in performing its gate-keeping inquiry in evaluating the reliability of expert testimony is the “known or potential rate of error.” *Nease v. Ford Motor Co.*, 848 F.3d 219, 229 (4th Cir. 2017) (citing *Daubert*, 509 U.S. at 594) (reversing district court’s decision to permit expert testimony, in part, because Fourth Circuit “could not assess the potential rate of error of [expert’s] methodology.”); *Hutchinson v. Hamlet*, No. C 02-

⁷ Understanding and Using ACS Data; What All Data Users Need to Know at 1, Chapter 1: Understanding the ACS: The Basics, U.S. Census Bureau, available at <https://www.census.gov/programs-surveys/acs/guidance/handbooks/general.html> (**Exhibit 5**) (emphasis added).

⁸ Understanding and Using ACS Data; What All Data Users Need to Know at 44, U.S. Census Bureau, Chapter 7: Understanding Error and Determining Statistical Significance (**Exhibit 6**).

⁹ A confidence *level* is different than a confidence *interval* or as it is also known – a margin of error. “The confidence level indicates the percentage of the time that intervals from repeated samples would cover the true value.” Reference Guide on Statistics, Reference Manual on Scientific Evidence 247 (3d ed. 2011). “Statements about confidence without mention of an interval are practically meaningless.” *Id.* at 248, 2011 WL 7724256.

974 JSW (PR), 2006 WL 1439784, at *4 (N.D. Cal. May 23, 2006) (granting *Daubert* motion because expert’s “methodology [does not] allow this Court to consider the “known rate of error or potential for error” associated with the methods [expert] employed for reaching them.”) (citing *Daubert*, 509 U.S. at 594). Professor Clark himself acknowledged that an MOE is necessary because “when statisticians and demographers make estimates, they recognize that there is some – because it is not a count – error in the result.” Ex. 2 at 26:19-22 (Clark Dep. Tr.). Without determining an MOE (or confidence interval) for his estimate of undocumented Hispanics in Census Tract 4406¹⁰ or CMS’s estimate of undocumented Hispanics in the Fairfax PUMA, Professor Clark’s statistical estimates of the percentage of undocumented Hispanics in the PUMA or the Tract are inherently unreliable.

Instead of determining an MOE for the undocumented Hispanic population, Professor Clark applied ACS’s MOE of 26% for ACS’s estimate of *all Hispanics* in Tract 4406 to conclude an interval of undocumented Hispanic population in Tract 4406 of 223-379. Ex. 1 at 5. This methodology is fundamentally flawed because the undocumented population is much more difficult to estimate, as Professor Clark himself concedes. But he still uses the ACS’s inapposite 26% MOE for ACS’s estimate of *all Hispanics* in Tract 4406 without increasing that MOE for the inherent difficulty in measuring the undocumented population. Ex. 1 at 5.¹¹ And, Professor Clark

¹⁰ CMS does not estimate the undocumented population at the tract level. Therefore, Professor Clark incorrectly applied CMS’s estimates for the larger sample size of the PUMA level to the tract level without increasing any MOE (CMS does not publish any below the national level) to adjust for the increased variability in a smaller sample size.

¹¹ Professor Clark should have determined a proper MOE for CMS’s estimates of the undocumented population at the county or PUMA level and his estimate of the undocumented population at the tract level by starting with the only known MOE that CMS publishes – CMS’s 9% national level MOE (Ex. 2 at 34:6-19) – and increasing the margin of error to adjust for the smaller sample sizes below the national level. As a sample size goes from a larger to a smaller sample, from the national to the state to the PUMA level, the MOE increases. Professor Clark testified that the Census Bureau follows the same concept. Ex. 2 at 32:9-13; *accord* Ex. 2 at 28:2-

never adopts ACS's inapposite MOEs for ACS's estimates of the total Hispanic population at any other level but the tract level.

Professor Clark therefore failed to determine any MOE for his estimate of the undocumented Hispanic population in the Tract (or even opine as to what the appropriate MOE would be for CMS's estimates at the larger PUMA level because CMS does not publish an MOE for those estimates).¹² Instead, he simply – and wrongly – applied an MOE for the estimate of a different population group that does not take into account the fact that the undocumented population is more difficult to estimate.¹³ Importantly, when pressed, Professor Clark admitted that “[p]erhaps the margin of error should be larger.” Ex. 2 at 78:12-13. How much so? The parties and the Court do not know. Of course, as noted, if an MOE is large, “the estimate may be seriously wrong”. Reference Guide on Statistics at 246.

Compounding these errors, Professor Clark never indicates what confidence level is being used for any of the estimates. Confidence levels and MOEs are interrelated. “The confidence level

7 (admitting that sample size affects reliability of an estimate and “larger samples are usually likely to be more accurate.”). Dr. Weinberg states that “[t]his is a basic statistical concept associated with sample surveys.” Weinberg Decl. ¶ 7 (Dkt. 248-3); Amicus Br. of Former Directors of the U.S. Census Bureau, *Evenwel v. Abbott*, 2015 WL 5675832, at *19 (U.S., 2015) (“The margin of error grows as the sample size decreases, so the smaller the area, the higher the possibility of error.”); see Understanding and Using ACS Data at 1 (“In general, the larger the sample, the smaller the level of sampling error.”) (Exhibit 5). Instead, of following the proper methodology, Professor Clark applied CMS's estimate for the amount of undocumented Hispanics in the PUMA area (31.4%) to the significantly smaller census tract without increasing the MOE and then adopted ACS's inapposite MOE for *his* estimate of undocumented Hispanics in Tract 4406.

¹² Likewise, Professor Clark never calculated an MOE for CMS's estimates of undocumented Hispanics at the County or State level. CMS does not publish an MOE for its undocumented estimates below the national level, and the margin of error (thus the rate of error) for these estimates is left unknown and undetermined by Professor Clark. Ex. 2 at 34:10-19.

¹³ Even then, Professor Clark only adopted the MOE from the ACS at the tract level and he never states that he is adopting ACS's MOE at any other level. Ex. 1 at 5 (“The estimate and variation for the Hispanic population in Tract 4406 is +/- 251.”).

indicates the percentage of the time that intervals from repeated samples would cover the true value.”¹⁴ According to the U.S. Census Bureau, an MOE “describes the precision of an ACS estimate at a given level of confidence. The confidence level associated with the MOE indicates the likelihood that the ACS sample estimate is within a certain range (the MOE) of the population value.”¹⁵ “The MOEs for published ACS estimates are provided at a 90 percent confidence level. From these MOEs, data users can easily calculate 90 percent confidence intervals that define a range expected to contain the true or population value of an estimate 90 percent of the time.”¹⁶

According to the Reference Guide on Statistics, “for a given confidence level, a narrower interval indicates a more precise estimate, whereas a broader interval indicates less precision. A high confidence level with a broad interval means very little, but a high confidence level for a small interval is impressive, indicating that the random error in the sample estimate is low.”¹⁷ “Statements about confidence without mention of an interval are practically meaningless[,]”¹⁸ but the reverse is also true, that mentions of intervals without statements about confidence are practically meaningless. This is because, as Professor Clark admitted, “you can have margins of error at various confidence levels.” Ex. 2 at 28:21-22; *id.* at 29:3-6 (“Demographers and census people use various levels of – various probability estimates ranging in the five -- .05 and .10, five and ten percent.”).¹⁹ A very low confidence level with a small interval means an estimate is simply

¹⁴ Reference Guide on Statistics at 247.

¹⁵ Understanding and Using ACS Data; What All Data Users Need to Know at 44, U.S. Census Bureau, Chapter 7: Understanding Error and Determining Statistical Significance (Exhibit 6).

¹⁶ *Id.*

¹⁷ Reference Guide on Statistics at 247-48.

¹⁸ *Id.* at 248.

¹⁹ Professor Clark then answers that another way of describing a confidence level is saying “90 percent” when answering if the Census Bureau uses a confidence level of 90 percent. Ex. 2 at 29:7-9.

not reliable because, logically, there is a very low confidence that after taking repeated samples of the population, the true value of the estimate would fall within the stated interval.

Professor Clark never provides any confidence levels regarding his estimates. But absent a confidence level, any MOEs for estimates of the undocumented population are rendered meaningless and wholly unreliable. Because Professor Clark's estimates do not enable this Court to consider the known or potential rate of error of those estimates, his opinions are not reliable and should be excluded. *See Crawford v. Newport News Indus. Corp.*, No. 4:14-CV-130, 2017 WL 3222547, at *7-8 (E.D. Va. July 28, 2017) (striking expert's conclusions because they relied on unreliable data); *Freeman*, 778 F.3d at 466 ("Expert testimony under Rule 702 is admissible if it 'rests on a reliable foundation and is relevant.'").

II. PROFESSOR CLARK'S FAILURE TO DETERMINE AN MOE FOR UNDOCUMENTED ASIANS LIKEWISE RENDERS HIS OPINION THAT UNDOCUMENTED HISPANICS ARE IMPACTED MORE THAN UNDOCUMENTED ASIANS COMPLETELY UNRELIABLE

As noted above, Professor Clark never determined an MOE for his estimate of undocumented Hispanics in the census tract and never determined an MOE for *CMS's* estimate of undocumented Hispanics in the Fairfax PUMA (or county or state level for that matter). He simply applied an inapposite 26% MOE from ACS's estimate of the total Hispanic population in the census tract and used that same MOE for his estimate of undocumented Hispanics in the census tract. This fundamental error alone renders Professor Clark's estimates completely unreliable. But it also undermines the reliability of his opinions, made for the first time in his rebuttal report, that undocumented Hispanics are impacted by the Policy more than undocumented Asians. *Anderson v. Westinghouse Savannah River Co.*, 406 F.3d 248, 263 (4th Cir. 2005) ("[I]n order to evaluate whether or not there is disparate impact in ratings, similarly situated persons who are being rated must be compared."). That opinion should be stricken as well.

Professor Clark failed to compare undocumented Hispanics to undocumented Asians in his opening report, and Dr. Weinberg pointed this out this flaw in Professor Clark's methodology in his opposing report. Exhibit 4 at 2, Opinion 5. Recognizing this problem, Professor Clark admitted in his reply report that he "did not adjust for the Asian undocumented population." Exhibit 3 at 2. Thereafter, in his rebuttal report, Professor Clark proceeded to analyze the undocumented Asian population in Tract 4406 using the "same methodological procedure" as in his original report. *Id.* at 2 § 3. He then adjusted his opinions dramatically to conclude that Latinos are not even twice as likely to be impacted as Asians, *id.*, even though he originally concluded that Hispanics are "nearly 7 times more likely to be undocumented than other groups". Ex. 1 at 5 § 5.²⁰

But in using his "same methodological procedure", Professor Clark never determined any

²⁰ In fact, the CMS PUMA data relied upon by Professor Clark does not demonstrate that Latinos are more likely to be impacted by a policy directed at illegal aliens than other similarly situated illegal-alien minority groups. The PUMA data relied upon by Professor Clark is set forth in Table B of exhibit 2 to Dr. Weinberg's declaration. Dkt. 248-3 at Ex. 3(2), table B. The CMS estimates at issue show the undocumented population for the relevant PUMA for 2012 and 2014. Weinberg Decl. ¶ 8 (Dkt. 248-3). For 2014, CMS estimated that Hispanics comprised 52.80% of the undocumented population of the PUMA. Dkt. 248-3 at Ex. 3(2), table B. The same CMS estimates of the undocumented population identifies Asians as comprising 37.10% of the undocumented population for the PUMA. *Id.* Even ignoring the MOE problems with these CMS figures (there are no MOEs), these CMS figures – which Professor Clark relied upon – thus demonstrate that in 2014 Latinos were, *at most*, only 1.4 times more likely to be impacted by a policy directed at illegal aliens as undocumented Asians. *See id.* (noting PUMA data of 52.80% Hispanic versus 37.10% Asian). These estimates are of course completely unreliable without any published margins of error.

More troubling about Professor Clark's opinions comparing the Asian and Hispanic population is that is that these same CMS estimates he relies upon demonstrate that, in 2012, Asians made up the majority of undocumented individuals for the relevant PUMA. Dkt. 248-3 at Ex. 3(2). This would mean in 2012, Professor Clark's opinions are flat wrong even using his flawed methodology, and it would Asians, not Hispanics, who would be purportedly impacted by the Policy. Professor Clark's opinions therefore are *ipse dixit* – not tethered to the data but cherry-picked from the years he believes support his opinion at the expense of the years that do not. *Freeman*, 778 F.3d at 469 (Agee, J. concurring) ("[C]ourts have consistently excluded expert testimony that "cherry-picks" relevant data.") (citing cases).

MOE for the undocumented Asian population. He simply assumed the ACS's MOE for the total Asian population in Tract 4406 and applied that to *his* estimate of the undocumented Asian population for Tract 4406 (because CMS does not publish estimates of the undocumented population at the tract level). In his expert report, Professor Clark states that the "estimate and variation for the Hispanic population in Tract 4406 is 957 +/-251, or 26%." Ex. 1 at 5. In his rebuttal report, Professor Clark states that the estimate of the total Asian population for Tract 4406 is "682 with a margin of error of +/-128." Exhibit 3 at 2. This is an MOE of 18.7%. Applying the 26% MOE to the estimates of Hispanic population in Tract 4406 and the 18.7% MOE to the estimates of the Asian population in Tract 4406 reveals that there is a significant range overlap where the Asian population is actually higher than the Hispanic population in Tract 4406.

Professor Clark then uses the ACS's inapposite MOE for the total Asian population and applies it to CMS's estimate for the undocumented Asian population at the PUMA level (16.5%), which he applies to estimate the undocumented Asian population in Tract 4406. Exhibit 3 at 2 § 3 ("Using the same logic as the estimation procedure for Hispanics, 16.5 percent of Asians in the PUMA are undocumented and if we apply that percent to the 682 Asians in Tract 4406, I estimate the undocumented Asian population in tract 4406 as 113 with a range of approximately 92-134."). But without Professor Clark determining an MOE for the undocumented Asian population, we do not know how reliable that estimate is – we simply know that 18.7% cannot be the accurate MOE because as Professor Clark admits, the undocumented population is more difficult to estimate. And, as noted, Professor Clark admitted that in regards to his methodology, "[p]erhaps the margin of error should be larger." Ex. 2 at 78:12-13.

Professor Clark's comparison between the impact on the Hispanic population and the impact on the Asian undocumented population in the Tract was already razor-thin (not even two

times). Without Professor Clark determining an MOE (or stating a confidence level) for these estimates of the undocumented population, including the Asian population, Professor Clark's comparison between these two populations is rendered wholly unreliable.

III. PROFESSOR CLARK INCORRECTLY APPLIED CMS'S 31.4% ESTIMATE FOR UNDOCUMENTED HISPANICS IN THE PUMA TO THE MUCH SMALLER CENSUS TRACT BASED ON HIS *IPSE DIXIT* ASSUMPTION THAT THE DEMOGRAPHICS WOULD BE THE SAME.

Professor Clark's methodology is also unreliable because he incorrectly applied CMS's 31.4% estimate of undocumented Hispanics in the relevant PUMA to the dramatically smaller census tract based on his *ipse dixit* assumption that the demographics would be the same in Census Tract 4406 as in the much larger PUMA. "As the Supreme Court has repeatedly held, 'nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert.'" *Cooper v. Smith & Nephew, Inc.*, 259 F.3d 194, 203 (4th Cir. 2001) (citing *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 157 (1999)).

Lacking *any* statistics for the undocumented population at the census tract level, Professor Clark incorrectly applied CMS's 31.4% estimate for the PUMA to the much smaller population of Tract 4406.²¹ Ex. 1 at 5 § 5. In doing so, Professor Clark assumed without any basis that the demographic makeup in the much smaller census tract would be the same as in the PUMA even though error increases with a smaller area. Amicus Br. of Former Directors of the U.S. Census Bureau, *Evenwel v. Abbott*, 2015 WL 5675832, at *19 (U.S., 2015) ("The margin of error grows as the sample size decreases, so the smaller the area, the higher the possibility of error.").

²¹ According to the statistics relied on by Professor Clark, Tract 4406 has a total population of 3,294 versus the much larger Fairfax PUMA's population of 157,949. Ex. 1 at 4 (second table noting Tract 4406's total population of 3,294); *id.* at 5 ("The PUMA has a population of 157,949.").

When questioned at his deposition as to why he assumed the demographic makeup would be the same in a much smaller sample (the census tract sample) as in the much larger PUMA such that the statistics would be the same for both, he admitted first that “[i]t is true that there’s variation across the PUMA” but claimed “there’s no evidence that the PUMA is so variable that it would make the estimating procedure improper.” Ex. 2 at 66:15-21. When pressed as to how he determined that there is “no evidence” of such variability, he stated that he “looked at the census tracts that made up the PUMA” and concluded that “[t]here are many of them like Tract 4406.” *Id.* at 66:22-67:2. He was then asked how he arrived at this conclusion: “Q: So is it your testimony that Tract 4406 is similar in its demographic make-up as other census tracts in the PUMA?” *Id.* at 67:3-6. He providing a telling response, stating: “***I did not do that analysis.***” *Id.* (emphasis added). Thus, Professor Clark admits he performed no analysis to determine if there was “no evidence” that the demographic makeup would be different in Tract 4406 than in the PUMA.

Accordingly, Professor Clark wrongly assumed – without any basis – that the CMS estimate of 31.4% undocumented Hispanics in the approximately 160,000 person Fairfax PUMA should apply to the approximately 3,300 persons in Census Tract 4406. His conclusion that 31.4% of the Hispanics in Tract 4406 are undocumented is thus connected to existing data only by Professor Clark’s *ipse dixit*, and his opinions should be excluded as unreliable.

IV. EXCLUSION IS ALSO WARRANTED UNDER RULE 403

Rule 403 separately counsels the exclusion of Professor Clark’s opinions, because their “probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.” Fed. R. Evid. 403. For the reasons stated above, Professor Clark’s opinions lack probative value. And, as noted above, Professor Clark dramatically adjusted his opinion in his rebuttal report to conclude Latinos are not even twice as likely to suffer an impact

from Defendants' Policy as Asians. This opinion is not probative on Plaintiffs' claims of disparate impact. Moreover, Professor Clark's opinions also are unfairly prejudicial. If a jury were to be told that Professor Clark—as an expert—has concluded that Latinos suffer a disparate impact from Defendants' Policy, the jury will give significant weight to that view. *See Daubert*, 509 U.S. at 595 (recognizing that “[e]xpert evidence can be both powerful and quite misleading,” and that the court’s role is to keep such misleading testimony from reaching the jury under Rules 702 and 403). The admission of his unfounded opinion testimony would do nothing but confuse and mislead the jury, and his opinions should also be excluded for this reason.

CONCLUSION

For the reasons stated above, Defendants respectfully request the Court strike the opinions and testimony of Professor William A.V. Clark and for such other relief as may be just and proper.

Respectfully submitted,

WAPLES MOBILE HOME PARK LIMITED
PARTNERSHIP, WAPLES PROJECT LIMITED
PARTNERSHIP AND
A. J. DWOSKIN & ASSOCIATES, INC.

/s/

Grayson P. Hanes (VSB No. 06614)
Michael S. Dingman (VSB No. 30031)
Justin deBettencourt (VSB No. 83806)
REED SMITH LLP
7900 Tysons One Place
Suite 500
McLean, Virginia 22102
(703) 641-4200 (Telephone)
(703) 641-4340 (Facsimile)
ghanes@reedsmith.com
mdingman@reedsmith.com
jdebettencourt@reedsmith.com
Counsel for Defendants

CERTIFICATE OF SERVICE

I hereby certify that on the 22nd day of September, 2020, I caused the foregoing to be filed electronically with the Clerk of the Court using CM/ECF, which will then send a notification of such filing to all counsel of record.

/s/

Grayson P. Hanes (VSB No. 06614)
Michael S. Dingman (VSB No. 30031)
Justin deBettencourt (VSB No. 83806)
REED SMITH LLP
7900 Tysons One Place
Suite 500
McLean, Virginia 22102
(703) 641-4200 (Telephone)
(703) 641-4340 (Facsimile)
ghanes@reedsmith.com
mdingman@reedsmith.com
jdebettencourt@reedsmith.com
Counsel for Defendants